

Amendments to the Specification:

Please replace the paragraph starting on page 5, line 15 with the paragraph below:

Even GAST is a good data structure that compactly represents strings, there are practical issues to use it for ANWE (Automated New Word Extraction). The space required are too large for constructing an efficient/feasible GAST from a large amount of corpus.